University of Michigan researchers for use in MEMs packaging technologies as a way to provide both mechanical support and thermal isolation to MEMS devices [95], and the technology was adapted for this project as a way to add support to the silicon rings and the center platform. The glass tether is shown in cross section in Figure 4.5. Pyrex 7704 glass was chosen because it is stiff, has a low thermal conductivity of about 1.15 W/m-K [96], and can be anodically bonded to silicon. The tether is designed to have a high aspect ratio. In this way, it can provide added stiffness and strength to the thermal structure without adding a significant amount of parasitic thermal conduction. A 2.5 mm tether with a 100 μm x100 μm cross sectional area will have a total thermal conductance of 6.96E-6 W/K, several orders of magnitude lower than the thermal conductances of the SiO₂ bridges that it reinforces.

**Figure 4.6:** 3-D Rendering FEM simulation of a glass tether, such as the one used in the silicon-glass-silicon process. The temperature difference was set at 10 °C.
4.3 Fabrication Process

This section will outline the fabrication process used to make the TE micro coolers discussed in this chapter. It will begin with an overview of the fabrication process presented in Table 4.1, Table 4.2, and Table 4.3. It will then discuss several of the challenges encountered in implementing this process, and the steps taken to mitigate those challenges. The section will finish with a detailed list of the fabrication steps used to complete the coolers presented in the following sections of this chapter. All processing steps described in the following sections were performed in the Lurie Nanofabrication Facility (LNF) at the University of Michigan.

Fabrication Process: Part 1 – Thermal Substrate

The fabrication process uses a wafer stack. The top layer implements the main thermal requirements necessary to make the TE micro-cooler, and the fabrication process for this wafer is presented below. It is referred to in the remainder of this chapter as the Thermal Substrate.
Table 4.1: Thermal Substrate fabrication process overview.

<table>
<thead>
<tr>
<th>Step 1.0</th>
<th>The process begins with a 100 μm thick silicon wafer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1.1</td>
<td>PECVD oxide is deposited on the front side of the wafer.</td>
</tr>
<tr>
<td>Step 1.2</td>
<td>Metal interconnects are deposited on top of the oxide layer.</td>
</tr>
<tr>
<td>Step 1.3</td>
<td>A shallow DRIE etch is performed on the backside of the wafer.</td>
</tr>
<tr>
<td>Step 1.4</td>
<td>A second DRIE etch is performed on the backside of the wafer, releasing the oxide bridges/membranes between isothermal silicon regions.</td>
</tr>
</tbody>
</table>

Fabrication Process: Part 2 – Structural Substrate

The remaining two wafers are used to add structural support to the Thermal Substrate. Collectively they are referred to as the Structural Substrate in the remainder of this chapter, and the fabrication process is summarized below in Table 4.2.
Table 4.2: Structural Substrate fabrication process overview.

<table>
<thead>
<tr>
<th>Step 2.0</th>
<th>This part of the process begins with a 500 µm thick silicon wafer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2.1</td>
<td>A shallow DRIE etch is performed on the front side of the wafer.</td>
</tr>
<tr>
<td>Step 2.2</td>
<td>A 100 µm thick glass wafer is bonded to the top of the silicon wafer, covering the recess.</td>
</tr>
<tr>
<td>Step 2.3</td>
<td>The glass is etched in concentrated HF solution to create a released glass tether.</td>
</tr>
</tbody>
</table>
Fabrication Process: Part 3 – Completing the Cooler

The cooler fabrication is completed by attaching the Thermal Substrate to the Structural Substrate, and depositing the thermoelectric materials as shown in Table 4.3 below.

**Table 4.3: Cooler completion process overview.**

| Step 3.1 | The thermal substrate is anodically bonded to the structural substrate to complete the final structure for the cooler. |
| Step 3.2 | The TE materials are deposited on top of the wafer stack. First Bi₂Te₃ is deposited, followed by Sb₂Te₃. |

**4.3.1 Processing Challenges**

**4.3.1.1 Wafer Selection**

Successfully executing the process outlined above required overcoming a number of fabrication challenges, many of which are rooted in the creation of the oxide bridges and the use of 100 μm silicon for the top layer of the cooler. A standard 500 μm wafer could not be used because of non-uniformity in the DRIE process step. The oxide
bridges were created by a thru-wafer DRIE etch, using the oxide on the wafer surface as an etch stop. However, the STS DRIE system available at the time in the LNF exhibited higher etch rates at the outer edges of the wafer than in the middle of the wafer. As a result, the etch reached the oxide at the edges of the wafer before reaching the oxide near the center of the wafer. The non-uniformity was high enough that the oxide in devices near the edge of the wafer was thinned and weakened before the etch was completed at the center of the device. By using a thinner 100 µm wafer, the etching process reached the oxide with less time delay across the wafer.

4.3.1.2 Wet Processing

Although the 100 µm thick wafer improved absolute etch uniformity, the use of such a thin substrate meant that several other standard fabrication process had to be modified to complete the process. The first of these included basic handling of the wafers, which were far more fragile than standard silicon substrates. The pressure exerted by turbulence in a normal DI water quench tank, or from agitating the 100 µm thick wafers in a chemical processing tank, was enough to fracture them. When chemical processing was being done, the 100 µm wafers needed to be placed in a cassette along with additional 500 µm thick wafers that alternated slots with the thin wafers. The thicker wafers prevented the water flow from exerting pressure on the thin wafers. The cassette containing the both thin and standard wafers could be placed in a spin-rinse drier (SRD) at the completion of a chemical processing step. This was preferable to hand drying the wafers because pressure exerted on the wafers by the nitrogen gun could also easily fracture them.
Processing steps that could not accommodate a cassette, such as gold etching and developing in AZ400K, required that only one wafer at a time be processed using an individual wafer holder. The wafer was then rinsed in the down-stream side of a cascade rinse tank with the nitrogen bubbler off, before being transferred to a cassette for a final rinse and drying in the SRD. Being able to handle the thin wafers and perform wet processing was an important accomplishment because up to 4 lithography steps needed to be done to the wafers before bonding them.

4.3.1.3 Oxide deposition

While successful wet chemical processing with the 100 µm wafers only required slight adjustments to normal procedures, the use of thin wafers had a larger impact on other areas of the fabrication process. The use of plasma enhanced chemical vapor deposition (PECVD) oxide deposited at 200 °C was a necessary choice, given the characteristics of 100 µm thick wafers. Oxide films deposited by thermal oxidation, low pressure chemical vapor deposition (LPCVD), and even PECVD at temperatures over 300 °C are denser and have higher mechanical strength than low temperature PECVD oxide. They also have higher levels of stress. Tests were done to use LPCVD oxide films and PECVD films deposited at 380 °C. However, for the anodic bonding that occurs later in the process to be successful, the backside of the wafer must be oxide free. With oxide on only one side of the wafer, the compressive stress in the film is high enough to visibly bow the wafer. It imparts enough curvature that the wafers cannot be clamped to the chuck of the photoresist spinner. The low temperature PECVD, on the other hand, has stress sufficiently low that wafer bow is not a hindrance to further process steps.
4.3.1.4 Oxide patterning and Metal Deposition

Patterning the oxide layer and depositing the interconnect metallization did not present a significant processing challenge beyond what was described above in relation to wet processing with the wafers.

4.3.1.5 Deep Reactive Ion Etching (DRIE)

Successfully performing DRIE on the 100 µm wafers required specialized processing. During the etch, the wafers are mechanically clamped to a chuck inside the tool, and helium gas flows over the backside of the wafer to aid with heat transfer from the wafer to the chuck. This prevents the photoresist on the wafer from burning, but it also puts pressure on the wafer that is high enough to break substrates with regions that are thinner than 250 µm. To prevent breakage, wafers that are etched deeply must be mounted on a carrier substrate. The normal mounting procedure is to spin a layer of 1827 photoresist (PR) on a carrier wafer. The device wafer is placed onto the layer of PR, and the two-wafer stack is cured in an oven at 110 °C for 30 min. However, if 100 µm thick wafers are mounted in this way and placed in the load lock of the DRIE tool, gas in the photoresist expands with enough force to shatter the 100 µm wafer. In order to process the thin wafers in DRIE, a special carrier wafer needed to be fabricated. The carrier had a pattern of criss-crossed channels 1 mm wide and 100 µm deep etched into a silicon wafer. This carrier was attached to the device wafer using the same PR bonding technique described above. When the photoresist is spun onto the carrier, the resist covers surface of the wafer and sides and bottom of the channels, but it does not fill the channels. However, when pumped to vacuum, the channels allowed the gasses to escape to the edges of the wafer and prevented the wafer from breaking.
An additional challenge associated with the DRIE steps was releasing the carrier from the device wafer without harming the structures. This was particularly true after the second DRIE step, when the oxide membranes were particularly fragile. Normally, wafers that have been bonded together with PR are released by soaking them in PRS2000, and then rinsing them in water. Placing the released wafers in flowing water caused many of the structures to break at the outer ring of oxide. To prevent this, the wafers were placed vertically in a beaker of acetone. After the acetone had dissolved the PR from between the wafers, the device wafer was carefully removed from the beaker and placed in a beaker of isopropyl alcohol to rinse off the acetone. Finally, the wafer was dried on a hot plate.

4.3.1.6 Wafer Bonding

Bonding also required a small change from the standard procedure. Normally, alignment of the wafers is performed using the Suss BA-6 bond aligner, but the alignment and chucking mechanism caused the thinned wafers to break. Instead, the wafers were manually aligned using through-hole alignment marks, and then placed in the bonding fixture.

4.3.1.7 Thermoelectric Material Integration

Integrating the thermoelectric materials with the structure described above was a significant challenge associated with this process for two reasons. First, the TE thin-films could not be patterned by lift-off or by etching. Lift-off was unavailable because of the high deposition temperature of the films. Most polymers, including the photoresists available in the LNF, will burn if heated to temperatures as high as 250 °C, where the thermoelectric deposition is optimized. When that happens, the patterns become
distorted and residue from the burned photoresist can contaminate the wafer and be difficult to remove. Etching could not be used in this case because of weak adhesion between the thin films and the substrates. Work by a previous University of Michigan student, Luciana da Silva suggested that the films etched unevenly and would quickly delaminate from the wafer [97].

Second, the Bi₂Te₃ and Sb₂Te₃ are contaminants for silicon processing, so an effort was made to place the TE deposition as the last step in the fabrication process. By that point the oxide bridges have already been released, so additional processing, particularly the wet processing needed for both etching and lift-off, would not be well tolerated. Instead shadow-masks are used to pattern the TE materials. The shadow-masks, shown in cross section in Figure 4.7, are made of silicon, and require a two-step process. First, large wide recesses are patterned on the backside of the wafer, and etched using DRIE to a depth of approximately ~400 µm. Then the front side of the wafer is masked with the pattern of the TE to be deposited, and it is etched through to the recesses. By thinning the wafer in the areas around the pattern that are being deposited, the shadow-mask interferes less with the flux of incoming material, and smaller, more precise patterns can be realized.
Figure 4.7: Illustration of a shadow mask cross sections (a) with a recess around the pattern, and (b) without a recess around the pattern. Using the recess allows flux impinging the wafer at shallower angles to reach the device wafer. This is important in systems where the source is not located along the axis perpendicular to the center of the wafer.

The shadow-mask has through-wafer holes that are inset with a cross pattern, which is used for alignment. On the device wafer there is a complimentary pattern. The two wafers are aligned manually under a microscope, and temporarily clamped together with spring clips. Stainless steel machine screws are then inserted into overlapping holes in the two wafers, and are used with a nut to clamp the wafers tightly during deposition. To prevent the shadow mask from exerting pressure on the fragile areas of the coolers, spacers were needed between the device wafer and the shadow mask. Pieces of 100 µm silicon were used, and were placed at the sites of clamping screws. Layers of aluminum foil, approximately 50 µm thick, were also used in some instances. Figure 4.8 shows a TE test pattern with high magnification views of specific features. Figure 4.9 shows top and bottom views of the shadow mask.
Figure 4.8: Images of TE material patterned by shadow masking. (a) The full test pattern. (b) The smallest element in the test pattern. It is 23 µm x 130 µm. (c) The largest element in the test pattern. It is 77 µm x 117 µm.

Figure 4.9 (a) View of the top side of a shadow mask showing the recess and the pattern that will be deposited on the cooler. (b) Pattern of the backside of the wafer showing the pattern that will be deposited. Taken from [75].
4.3.2 Detailed Fabrication Process

4.3.2.1 Thermal Substrate

1. A 1.5 µm thick layer of SiO$_2$ is deposited in the GSI PECVD tool with a substrate platen temperature of 200 °C.

2. A layer of 1827 photoresist is spun onto the wafers at 1000 rpm for 300 seconds, and the wafers are baked on a hotplate at 110 °C for 90 seconds.

3. The wafers are exposed on the MA-6 aligner for 20 seconds. They are developed in MIF 319 for 80 seconds, rinsed for 3 minutes in DI water, and dried in an SRD.

4. 400 Å of Cr and 5000 Å of Au are deposited in the Enerjet e-beam evaporator.

5. Lift off is performed in acetone with the aid of ultra-sonic agitation. The wafers are rinsed in isopropyl alcohol and dried on a hot plate.

6. A layer of 9260 is spun on the backside of the wafer at 2000 rpm for 30 seconds, and baked in an oven for 20 minutes at 90 °C.

7. The photoresist is exposed on the MA-6 for 55 seconds. It is developed in AZ400K for 90 seconds, rinsed for 3 minutes in DI water, and dried in an SRD.

8. A layer of 1827 is spun onto the backing wafer at 3000 rpm for 30 seconds. The Thermal Substrate is mounted to the spun photoresist, and both wafers are baked at 110 °C for 20 minutes.

9. The wafer is etched in the STS DRIE tool for 10 minutes, using the GC standard recipe. This results in an etch that is ~30 mm deep.
10. The Thermal Substrate is released from the backing wafer in Acetone. It is rinsed in isopropyl alcohol, and dried on a hot plate.

11. A layer of 9260 is spun on the backside of the wafer at 2000 rpm for 30 seconds, and baked in an oven for 20 minutes at 90 °C.

12. The photoresist is exposed on the MA-6 for 55 seconds. It is developed in AZ400K for 90 seconds, rinsed for 3 minutes in DI water, and dried in an SRD.

13. A layer of 1827 is spun onto the backing wafer at 3000 rpm for 30 seconds. The Thermal Substrate is mounted to the spun photoresist, and both wafers are baked at 110 °C for 20 minutes.

14. The wafer is etched in the STS DRIE tool for 22 minutes, using the GC standard recipe. This etches through the wafer and stops on the oxide layer.

15. The Thermal Substrate is released from the backing wafer in Acetone. It is rinsed in isopropyl alcohol, and dried on a hot plate.

16. Both the front and the back sides of the wafer are cleaned using O2 plasma.

### 4.3.2.2 Mechanical Supporting Substrate

1. A layer of 9260 is spun on the backside of the wafer at 2000 rpm for 30 seconds, and baked in an oven for 20 minutes at 90 °C.

2. The photoresist is exposed on the MA-6 for 55 seconds. It is developed in AZ400K for 90 seconds, rinsed for 3 minutes in DI water, and dried in an SRD.
3. The wafer is etched in the STS DRIE tool for 12 minutes, using the GC standard recipe.

4. The photoresist is stripped using PRS 2000. The wafer is rinsed for 3 minutes in DI water and dried using an SRD.

5. A 100 µm thick pyrex glass wafer is anodically bonded to the top side of the silicon wafer using the SB6e. The bond is performed at 450 °C and 400 VDC.

6. The wafer is cleaned using a 1:1 mixture of H₂O₂ and H₂SO₄ (piranha) for 10 minutes. It is rinsed for 3 minutes in DI water and dried using an SRD.

7. The wafer is etched in buffered hydrofluoric acid (BHF) for 5 minutes. It is rinsed for 3 minutes in DI water and dried using an SRD.

8. 1000 Å of Cr and 5000 Å of Au are deposited in the Enerjet e-beam evaporator.

9. A layer of 9260 is spun on the backside of the wafer at 2000 rpm for 30 seconds, and baked in an oven for 20 minutes at 90 °C.

10. The photoresist is exposed on the MA-6 for 55 seconds. It is developed in AZ400K for 90 seconds, rinsed for 3 minutes in DI water, and dried in an SRD.

11. The Au and Cr layers are etched using wet etchants with a 3 minute rinse in DI water following each etch. The wafer is dried in an SRD.

12. The glass layer is etched though using concentrated HF solution for 18 minutes.
13. The photoresist is stripped using PRS 2000. The wafer is rinsed for 3 minutes in DI water and dried using an SRD.

14. The Au and Cr layers are stripped with a 3 minute rinse in DI water following each etch. The wafer is dried in an SRD.

4.3.2.3 Wafer completion

1. The thermal substrate and structural substrate are manually aligned. A wafer with a thin layer of oxide is placed on top of the Thermal Substrate to prevent intermetallics of gold from forming with the bond chuck during bonding. The Thermal and Structural substrates are bonded at 350 °C and 400 VDC.

2. The shadow mask for the Bi₂Te₃ pattern is aligned to the device wafer and clamped. The wafer is loaded into the TE evaporator and pumped to a pressure of less than 2E-6 Torr. Bi₂Te₃ is deposited using a flux ratio of Te:Bi of 3.0:1, and substrate temperature of 260 °C. After deposition the wafer is allowed to cool to less than 100 °C before venting the chamber.

3. The shadow mask for the Sb₂Te₃ pattern is aligned to the device wafer and clamped. The wafer is loaded into the TE evaporator and pumped to a pressure of less than 2E-6 Torr. Sb₂Te₃ is deposited using a flux ratio of Te:Sb of 2.4:1, and substrate temperature of 230 °C. After deposition the wafer is allowed to cool to less than 100 °C before venting the chamber.
4.4  5-Stage Coolers

4.4.1  Specifications

Initial devices fabricated using the process above were designed to validate the process and test the thermal resistance of the structures. They were designed secondarily to demonstrate a functioning multi-stage micro-scale cooler. However, the cooler design used in these test structures was not rigorously optimized. The initial devices all had 5 stages arranged in concentric circles as shown in Figure 4.10. The thermocouples were 30 \(\mu\)m long and 100 \(\mu\)m wide. The first stage had 36 TC, the second stage had 20, and the third, fourth and fifth stages had 4, 2, and 1 TCs, respectively. The stages were arranged in concentric circles with stage 1 having a diameter of 6.2 mm. The full design parameters are summarized in Table 4.4. The 1\textsuperscript{st} and 2\textsuperscript{nd} stages could both be driven by individual currents, and the inner three stages could be driven by a third input current. This combination of oxide bridges and metallization will be referred to as M3. Four variations of the glass supporting structure were tested. The first variation (T1) bonded only to the center of the cooler. The second variation (T2) supported the center and silicon between the 2\textsuperscript{nd} and 3\textsuperscript{rd} stages. The third variation (T3) supported the silicon between the 1\textsuperscript{st} and 2\textsuperscript{nd} stages, between the 2\textsuperscript{nd} and 3\textsuperscript{rd} stages and at the center, and the 4\textsuperscript{th} variation (T4) had no supporting tether at all. T1, T2 and T3 are shown in Figure 4.11 and Figure 4.12 shows an SEM of the T2 tether design.
Table 4.4: Critical dimensions of the 5-Stage thermoelectric cooler

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter of Stage</td>
<td>6.2mm</td>
<td>4.1mm</td>
<td>2.1mm</td>
<td>1.8mm</td>
<td>1.5mm</td>
</tr>
<tr>
<td>Number of TCs</td>
<td>36</td>
<td>20</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TE Width</th>
<th>TE Length</th>
<th>TE Thickness</th>
<th>TE Contact Area</th>
<th>Oxide Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>80µm</td>
<td>30 µm</td>
<td>2µm</td>
<td>50µm x 80µm</td>
<td>1.5µm</td>
</tr>
</tbody>
</table>

Figure 4.10: (a) A top view of a 5-stage cooler built using the Silicon-Glass-Silicon process. (b) The inner three stages of the same 5-stage cooler, showing the temperature sensor.
Figure 4.11: Pictures of (a) the T1 tether, (b) the T2 tether, and (c) the T3 tether designs. The pictures have been artificially colored to increase the contrast to the background and to illustrate different regions of the tether designs. Red areas indicate the regions that will be bonded to the Thermal Substrate. Blue regions are unbonded.

Figure 4.12: An SEM image of the T2 tether design used with a 5-Stage cooler. This shows the tether suspended over of 10 µm recess in a silicon wafer.

Additionally, 3 sets of test structures were placed on the wafer simply to test the thermal resistance of the cooler structure with as little metal as possible present. Two sets of these devices were structurally identical to the cooler structures except for the metallization. One set implemented only a single resistive temperature sensor at the center of device. The combination of metal interconnects and oxide bridges will be
referred as design M1. A second set of devices contained resistive temperature sensors on the silicon between the 1st and 2nd stages, between the 2nd and 3rd stages, and at the center of the device. This combination of metal interconnects and oxide bridges will be referred as design M2. Variations of both these sets included devices with all 4 glass tether designs. A third set of devices with 60μm long oxide bridges were implemented for testing thermal isolation. These all used the T1 tether design. This structure was used to test the upper limits of thermal resistance that could be achieved with this fabrication process. The combination of metal interconnects and oxide bridges will be referred to as design M4. M1, M2, and M3 are shown in Figure 4.13

![Figure 4.13: Top views of three metal designs, showing (a) M1, (b) M2, and (c) M3. The metallization used in M1 is also used in the M4 pattern (not shown) but the oxide bridges are longer in the M4 design.](image)

4.4.2 Results and Discussion

The structure was first tested to verify the thermal isolation between the center of the device and the substrate. The measurement was taken by operating the thermistor in a four-point configuration, as illustrated in Figure 4.14. The current through and voltage across the resistor were recorded, and resistance and power dissipation were calculated. Then the input current was increased and the measurements were repeated. The
increased power dissipated in the resistor resulted in an increase in temperature at the center of the device. The temperature change was sensed as an increase in the calculated resistance of the thermistor. By using the thermal co-efficient of resistance for gold, the observed change in resistance was used to calculate the change in temperature.

Measurements were taken at several different current levels, and the relationship between input power and temperature was analyzed using a linear regression. The slope of the line, with units of Kelvin/Watt, corresponds to the thermal resistance of the completed structure. Measurements were first performed at atmospheric pressure on devices that did not yet have the thermoelectric materials. The results showed thermal isolation values between 300 K/W and 600 K/W. Then the wafer was placed in a vacuum probe station and tested at pressures of less than 10 mTorr. At low pressure the devices showed thermal resistances between 8000 K/W and 16300 K/W. The wide range in values is a result of variations in the design described above. An example of data taken at atmosphere and under vacuum is shown in Figure 4.15 and a summary of the cross-wafer performance, indicating the various designs used, is given in Figure 4.16. This result also shows the importance of operating at vacuum. At atmosphere, the parasitic thermal conduction due to air is an order of magnitude greater than the thermal conduction of the TE. In Chapter 3, it was shown that when this is the case, the expected cooling is only a small fraction of what the TE material is capable of achieving.
Figure 4.14: A schematic representation of the thermal isolation measurement circuit. The four-point measurement system is used to avoid measuring changes in resistance due to parasitic elements that are not located at the center of the device.

Figure 4.15: Thermal resistance measurement of the 5-stage cooler in air and in vacuum. The thermal resistance in vacuum is 27 times higher than the thermal resistance in air.
Figure 4.16: A wafer map showing thermal resistance measurements across several variations of cooler designs. Die in green were testable. The top number indicates thermal resistance in K/W, and V-x indicates a variation designator as explained in the text.

A modified temperature measurement scheme was used when evaluating the performance of the coolers. As a result of limited access to the wafer in the vacuum chamber, a four-point measurement could not be used at the time these wafers were tested. Instead a two-point measurement was used on the thermistor. This resulted in some measurement uncertainty because the resistance being measured between the contacts was not undergoing a uniform temperature change. To place maximum and minimum bounds on the temperature change, observed changes in resistance were converted to temperature changes in two ways. First, it was assumed that change in resistance was due to a uniform cooling of the entire structure, and resulted in the minimum bound for the actual cooling. Next, it was assumed that only the center of the device was cooled, and only the resistance at the center of the device contributed to the
observed change in resistance. This resulted in the maximum bound on the observed cooling. Reported data is the average of the two bounds. Using this measurement scheme, cooling performance was evaluated using two different configurations of the cooling circuit. In the first configuration, all five stages of the cooler were connected in series. The devices were tested under vacuum, and a maximum temperature between the device center and the substrate of 3.45 °K was achieved. The average cooling for all four tested devices was 3.03 °K degrees.

In the second configuration, only the outer two stages of the device were used, again connected in series, and this resulted in a maximum achieved cooling of 3.8 °K. In this configuration, the average cooling was 3.26 °K. Later fabrications of the same cooler design produced better results and were able to utilize the 4 point measurement scheme for temperature sensing. The best of these coolers achieved 8K of cooling, and matched well with the performance predicted by the model, as shown below in Figure 4.17.
Figure 4.17: Measured performance of a full 5-stage cooler plotted versus input current. Power consumption is plotted using the right hand axis, and modeled cooling performance is plotted using nominal (based on the properties reported in Chapter 2) and fitted material properties.

4.5 6-Stage Coolers

To increase performance beyond the initial 5-Stage device new coolers were designed. The layout of these devices was informed by the 1-D Mathematica and Matlab based models described in Chapter 3. Both the number of stages and the distribution of thermocouples were changed from the previous version. The new version of the cooler included 6 stages instead of 5, and the ratio of thermocouples in consecutive stages was changed to be 2 for all stages, instead of varying as in the five stage cooler. By decreasing the spacing between TCs, the overall size of the cooler was not increased, despite the addition of the sixth stage. The geometry of the TE elements was also changed slightly. The widths of the Bi$_2$Te$_3$ and Sb$_2$Te$_3$ was changed to compensate for the difference in Seebeck coefficient and resistance. The Bi$_2$Te$_3$ was 100 µm wide while the
Sb$_2$Te$_3$ was 120µm wide. The length of the TE element was kept constant at 30 µm, but the contact area were increased to 100 µm x 100 µm. The properties are summarized in Table 4.5, and a top-down image of the cooler is shown in Figure 4.18.

![Figure 4.18: Top view of the 6-Stage cooler, Version 1. It shows the thermoelectric elements, the temperature sensor, and the silicon rings.](image)

### Table 4.5: Critical dimensions of the Version-1 6-Stage thermoelectric cooler

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage width</td>
<td>5.7 mm</td>
<td>3.5 mm</td>
<td>3.0 mm</td>
<td>2.5 mm</td>
<td>2.0 mm</td>
<td>1.5 mm</td>
</tr>
<tr>
<td>Number of TCs</td>
<td>64</td>
<td>32</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TE Width</th>
<th>TE Length</th>
<th>TE Thickness</th>
<th>TE contact area</th>
<th>Oxide Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sb$_2$Te$_3$: 120 µm</td>
<td>30 µm</td>
<td>Sb$_2$Te$_3$: 3.8 µm</td>
<td>100 µm x 100 µm</td>
<td>1.8µm</td>
</tr>
<tr>
<td>Bi$_2$Te$_3$: 100 µm</td>
<td></td>
<td>Bi$_2$Te$_3$: 2.9 µm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The glass tether structure was also changed for this revision of the cooler. Instead of long wrapped coil structures, such as the one presented in Figure 4.12, the new devices had a straight tether which spanned from corner-to-corner of the device as shown in Figure 4.19. The tether is still made using 100 µm thick glass, and is designed to have an
average width of 100 µm, but the length is significantly shorter than in the previous
designs. Each half of the tether is 3 mm long. Assuming a thermal conductivity for glass
of 1.15 W/m-K, this results in a total thermal conductivity of 7.7E-6 W/K. This is still an
order of magnitude less than the thermal conductivity presented by the oxide bridges,
which for this design present a total thermal conductance of 8.5E-5 W/K. It therefore has
only a mild influence on the cooling ability of the device. For example without a tether
the device above is modeled as being able achieve ΔT=23.8 K. With the tether the cooler
is modeled as reaching ΔT=22.8 K, less than a 5% decrease in performance.

![Figure 4.19: An image of the tether used for Version 2 of the 6-stage cooler. The picture has been artificially colored to increase the contrast to the background.](image)

4.5.1 Results and analysis

As before, the coolers included a Cr/Au resistive temperature sensor that was used
to measure the cooling achieved by the device, and the device was tested in a Lakeshore
cryonics vacuum probe station. The fabricated coolers were tested with input current varying from 0 mA to 6.5 mA. The resistive temperature sensor dissipated 100 µW of power during the testing. This amount of power dissipation by the temperature sensor raises the temperature of the cold stage above ambient by less than 1 K. Figure 4.20 shows the measured data for a 6-stage cooler. The maximum temperature difference observed was 16.6 K, and occurred with a current input of 5.1 mA. The total resistance of the cooler was 1071 Ω, resulting in a total power consumption of 26 mW.

Additional analysis of the data was performed using the Mathematica-based analytical model described in Chapter 3. To estimate the effective material properties that had been achieved in-situ, the model was fitted to the collected data using the material properties as fitting parameters. The fit was performed as if both deposited TE materials had the same properties. Contact resistance was first used to adjust the total resistance of the simulated cooler to the measured resistance of the actual cooler, assuming a resistivity of 20 µΩ-m. (The same results can be achieved by assuming contact resistivity as 0 Ω-m, and varying the resistivity of the thin films to match the total resistance of the completed cooler). Next, the Seebeck coefficient was adjusted so that the measured optimum current matched the simulated value, given the measured resistance of the cooler. Finally the thermal conductivity of the TE material was adjusted to fit the magnitude of the maximum simulated ΔT with the measured results. The model used included the SiO₂ membrane, which was assumed to have a thermal conductivity of 1.15 W/m-K. For this version of the 6-stage cooler, Version 1, the fitting process estimated the Seebeck coefficient to be 180 µV/K, and the thermal conductivity to be 1 W/m-K.
4.5.2 Modified 6-Stage Cooler (Version 2)

The 6-stage cooler was updated with a number of design modifications to enhance cooler performance. Most significantly, the size and shape of the TE-metal contact was changed in an effort to reduce contact resistance. To do this, several changes were made to the contact regions. For TC with a large spatial separation between adjacent stages, such as the space between stage 1 and stage 2, or stage 2 and stage 3, the size of the contact region was increased. To determine the best shape for these contacts, 3-D finite element simulations were performed using a variety of geometries. These results are plotted in Figure 4.21. They show that any increase in the size of the contact region is helpful, but increasing the width of the contact region provides a greater benefit than increasing the length. This is particularly true for contacts longer than 100 μm. Where possible, contacts with a 200 μm width and a 100 μm length were used. However, at stage 1 the perimeter was confined by overall size constraints, and there was not enough
room to add such wide contacts. In this case, the contacts were set to be 100 \( \mu \text{m} \) by 200 \( \mu \text{m} \) long. To accommodate all of the thermocouples in the 2\(^{nd}\) stage using the new, larger contact scheme, the perimeter needed to be increased. The length of one side of the 2\(^{nd}\) stage was 4.23 mm, compared to 3.5 mm in Version 1.

![Graph of simulated TE-metal contact resistance for a variety of contact geometries. The assumed TE resistivity for these simulations is 20 \( \mu \Omega \cdot \text{m} \), and the contact resistivity is 1E-8 \( \Omega \cdot \text{m}^2 \). The current enters the contact in the TE layer and leaves in the metal layer.](image)

**Figure 4.21:** Simulated TE-metal contact resistance for a variety of contact geometries. The assumed TE resistivity for these simulations is 20 \( \mu \Omega \cdot \text{m} \), and the contact resistivity is 1E-8 \( \Omega \cdot \text{m}^2 \). The current enters the contact in the TE layer and leaves in the metal layer.

In addition, it was observed that the total number of contacts at the inner stages could be reduced by passing individual TE elements across multiple stages, while at the same time reducing the size of the silicon rings acting as thermal conductors. This concept is illustrated in Figure 4.22. The contact regions are still deposited over gold as before, and the size of these contacts was also informed by FEM simulations. Contacts that are 200 \( \mu \text{m} \) and 100 \( \mu \text{m} \) long were chosen. The results of those simulations are shown in Figure 4.23. The purpose for these modifications was to reduce the effect of contact resistivity on the performance of the cooler. A cooler fabricated with this process had a resistance of 854 \( \Omega \), compared to 1071 \( \Omega \) for version 1. A top view of the revised 6-stage cooler is shown in Figure 4.24. In addition to the changes made regarding the
contacts, the entire surface of the inter-stage regions were coated with gold to minimize any potential radiation losses.

Figure 4.22: Comparison of contacts in Version 1 (right side) and Version 2 (left side). The images show contact dimensions and the current path through 3 stages of each cooler.
Figure 4.23: Simulated resistance of a TE element deposited over gold. The TE resistivity is assumed to be 20 $\mu\Omega\cdot\text{m}$, and the contact resistivity is $1\times10^{-8} \Omega\cdot\text{m}^2$. The current enters and leaves the element through the TE material.

Table 4.6: Critical dimensions of the Version-2 6-Stage thermoelectric cooler

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage Width</td>
<td>5.4 mm</td>
<td>4.23 mm</td>
<td>2.28 mm</td>
<td>2.02 mm</td>
<td>1.76 mm</td>
<td>1.5 mm</td>
</tr>
<tr>
<td>Number of TCs</td>
<td>64</td>
<td>32</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TE Width</th>
<th>TE Length</th>
<th>TE Thickness</th>
<th>TE Contact Area</th>
<th>Oxide Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Sb_2Te_3$: 100 µm</td>
<td>30 µm</td>
<td>$Sb_2Te_3$: 3.2 µm</td>
<td>200 µm x 100 µm</td>
<td>1.5 µm</td>
</tr>
<tr>
<td>$Bi_2Te_3$: 100 µm</td>
<td></td>
<td>$Bi_2Te_3$: 2.7 µm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.5.3 Test Results

Testing was again performed under vacuum using a Lakeshore Cryogenics vacuum probe station. The temperature at the center of the cooler was measured using a resistive thermal temperature sensor fabricated in the chrome/gold interconnect layer. A source/meter with four point sensing capability was used to supply 100 uW of power to the sensor and to measure the voltage across the sensing element. Current was supplied to the thermocouples by a second source/meter, and an automated test program was used to test the cooler at a variety of input currents. Figure 4.25 shows the current vs. temperature relationship of several devices, with the best performing device achieving a temperature differential of 22.2 K at an input current of 5.5 mA. This corresponds to a total power consumption of 24.7 mW.
The Mathematica-based analytical model was fitted to this version of the cooler in the same way that it was fitted to version 1 of the 6-stage cooler. The fitted curve shown with the measured data in Figure 4.26 indicates a thermal conductivity for the TE material of 1 W/m-K, but the Seebeck is lower at 166 µV/K. This physical variation in TE material between the coolers is likely due to inaccurate control of material flux rates and substrate temperatures inside the deposition chamber.

To understand how the structure of the cooler affected the overall results, the ideal model was fitted to the data and used to calculate an effective Z. This analysis answers the question: In the absence of parasitic thermal conductance and parasitic resistance, what intrinsic figure of merit would produce a 6-stage cooler with the measured performance curve? The performance of Version 1 would be replicated by using a material with a Z of 1.6e-4 K⁻¹, corresponding to a ZT of 0.048 at 300 K. Version 2 produced a slightly better result, requiring material with a Z of 2.2e-4 K⁻¹, or a ZT of 0.066 at 300 K. However, both are low compared to the intrinsic material properties described in Chapter 2. This indicates that the TE material is not being effectively utilized. As already discussed, this is most likely due to the excess oxide used in the inner stages of the coolers.
Figure 4.25: Measured performance of several Version 2, 6-stage coolers plotted versus input current. All demonstrate cooling of greater than 19 K.

Figure 4.26: Measured performance of Version 2 of a 6-stage cooler, plotted together with simulated performance that has been fit to measured data to extract the material properties of the TE.

4.6 1-Stage cooler

A 1-stage cooler was also fabricated on the same layout with version 2 of the 6-stage cooler. The legs of the TCs were 60 μm long by 100 μm wide, and the contacts
were 100 µm x 100 µm. There were 16 total TCs, and the width of the stage was 1.5 mm. A top-down view of the 1-stage cooler is shown in Figure 4.27. The 1-stage cooler was fabricated on the same substrate as Version 2 of the 6-Stage cooler, and was tested using the same procedure as the previous 6-stage coolers. A plot of cooler performance versus input current is shown in Figure 4.28. The 1-stage cooler was able to achieve a temperature differential of 17.9 K at a current input of 8.5 mA.

Figure 4.27: Top view of a single stage cooler with 16 TCs.
Using the same curve fitting technique described above, the average Seebeck coefficient was found to be 185 $\mu$V/K, again with an estimated thermal conductivity of 1 W/m-K. The effective $Z$ was also calculated for the 1-stage cooler and found to be $2.6 \times 10^{-4}$ K$^{-1}$, corresponding to a $ZT$ of 0.14. While significantly closer to the measured intrinsic value, this low number indicates that the performance of the 1-stage cooler is also hampered by parasitic effects due to contact resistance and the presence of the oxide membrane. This Seebeck coefficient is larger than the value found for the Version 2 6-stage cooler described above, even though both were fabricated on the same wafer. However, it has previously been shown that the Seebeck coefficient of the coevaporated thin films is closely related to the substrated temperature during deposition [74]. In the case of the 6-stage cooler, the center stages are highly isolated from the general substrate and thus the temperature of inner stages of the cooler may vary from the temperature of

![Figure 4.28: Measured performance of a 1-stage cooler, plotted together with simulated performance that has been fit to measured data to extract the material properties of the TE.](image)
the substrate during deposition. This likely degrades the quality of the material at the inner stages.

### 4.6.1.1 Thermally Loaded Testing

Both the 1-stage cooler and version 2 of the 6-stage cooler were tested under thermal loads. A resistive temperature sensor supplied loads of varying magnitude. As the thermal load was raised, both coolers were able to generate higher differences in temperature between their off states and their optimum current point. In addition, both the coolers exhibited an increase in the magnitude of the optimal current. With the cooler off, a 50 mW load was applied, raising the center of the 1-stage device to 450.3 K. When a current was applied, \( I_{\text{opt}} \) was found to have increased to 12 mA, and the center region was cooled to 415.0 K. This represents a change of 35.3 K, compared to only 18.0 K using an 8 mA current when the device was loaded with 100 uW. Similarly with a 25 mW load, the 6-stage device was able to reduce the temperature of the cooler region from 494.9 K to 465.8 K, a change of 29.1 K with an optimum current of 7 mA. This apparent increase in performance is a result of the temperature dependence of \( \Delta T_{\text{max}} \) as described by Eqn. 3.9. The data from the 1-stage cooler is shown in Figure 4.29 and Figure 4.30 along with curves fitted to the ideal cooling equations. The change in temperature of the center is fitted to Eqn. 3.9 and results in a calculated \( Z=4.3E^{-4}\pm2.6E^{-5} \) K\(^{-1} \) with a 95% confidence interval. This corresponds to a ZT of 0.13±0.007, which is similar to the ZT value calculated above for the 1-Stage cooler. The relationship between temperature and the optimum current was also fitted to Eqn. 3.7, and produced \( \alpha/R_c=2.87E^{-5}\pm5E^{-7} \), and fits the data with a sum of squares error of 5.56E-8. These curves fit the data well, and provide a reasonable estimate of the figure of merit where compared to the earlier
calculations based on the cooling vs current curves. The data for the 6-stage cooler is presented in Figure 4.31 and Figure 4.32. However, because the model of a 6-stage cooler is much more complex, it is difficult to extract a value of effective $Z$ from this data.

The lower loads were used to measure the thermal resistance of the total structure. A 1-stage cooler increased 2.7 K with a 1 mW load, leading to a calculated thermal resistance of 2700 K/W. This indicates that with a 6.6 mW load, and using the optimum input current, $\Delta T=0$. At this operating point the cooler uses 14.6 mW of power giving it an approximate COP of 0.45. The Version 2, 6-Stage cooler that was tested with thermal loads increased 8.9 K for a base temperature of 379.2 with a 1 mW load. It used 20.8 mW of power. This corresponds to 8900K/W thermal resistance, 2.3 mW of load at $\Delta T=0$, and an approximate COP of 0.1.

![Figure 4.29: A 1-stage cooler tested at various thermal loads. The temperature change achieved and the optimum current both increase as the input load increases.](image)
Figure 4.30: Variation in $\Delta T$ and $I_{opt}$ with respect to cold stage temperature for a 1-stage cooler. The ambient temperature is maintained at 300 K, and the temperature of the center region is increased by applying a thermal load with a resistive element.

Figure 4.31: A Version 2 6-stage cooler tested at various thermal loads. The temperature change achieved and the optimum current both increase as the input load increases.
Figure 4.32: Variation in $\Delta T$ and $I_{\text{opt}}$ with respect to cold stage temperature for a 6-stage cooler. The ambient temperature is maintained at 300 K, and the temperature of the center region is increased by applying a thermal load with a resistive element.

4.7 Process and Design Shortcomings

4.7.1 Parasitic Thermal Conduction

Although the process detailed in the beginning of this chapter allowed for the fabrication of coolers with moderate temperature differentials and at low power consumption levels, it still presented shortcomings that affected both performance and usefulness in an integrated system. First, the oxide bridges covered nearly the entire perimeter of each stage. Although the oxide was thinner than the TE, this arrangement meant that the total cross section of oxide spanning each stage was nearly as large and in some cases much larger than the cross-section of the TE. For example, consider the 6th stage in the device above. It has two TCs that are 30 $\mu$m x 100 $\mu$m x 2.5 $\mu$m. They span
a bridge that is 30 µm long and has a perimeter of 6 mm and a thickness of 1.5 µm. The TCs will have total thermal conductance of 3.8e-5 W/K, while the oxide will have a thermal conductance of 3.4e-4 W/K. Remembering from Chapter 3 that the effective figure of merit (\( Z_{\text{eff}} \)) is proportional to \( K_{\text{TE}}/K_{\text{par}} \), this calculation suggests that \( Z_{\text{eff}} \) at the inner stages will be at best only 11% of intrinsic value. This suggests that there is very little cooling occurring at the inner stages, and this explains part of why there is so little separation in performance between the 1-stage cooler and the 6-stage cooler. If the oxide is only present below the TE elements of both devices, modeling shows that the 1-stage device would achieve a temperature differential of 23 K, and the 6-stage device would achieve a temperature differential of 31.5 K.

### 4.7.2 Contact Resistance

The second issue is contact resistance. Although steps were taken to mitigate contact resistance, it was still present and it diminished the ability of the cooler to perform. The contacts used in this research were all made from gold. It is possible that other materials would offer lower contact resistances; however, we were unable to clean the surface of the wafers in-situ and consequently had to choose a metal that does not oxidize, such as gold. An additional strategy, which will be employed in later designs shown in chapter 5, is to mitigate the effects of the contact resistance without reducing it. This can be done by increasing the aspect ratios of the TE elements in order to increase their resistance. As \( R_{\text{TE}} \) grows much larger than \( R_{\text{par}} \) the relative effect of \( R_{\text{par}} \) decreases, even though its absolute value remains the same.
4.7.3 Die Attachment and support

Finally, this generation of devices did not have the necessary structural rigidity to allow for die bonding operations, which would be necessary for the device to be integrated with other MEMS and electronics. The Thermal Substrate is released very early in the overall process, and the method of release, backside DRIE, means that the die could not be bonded prior to this step. Once the Thermal Substrate is integrated with the Structural Substrate it gains some support from the glass tethers. This is enough to prevent the coolers from breaking during handling and testing; however, it is not enough to prevent the oxide bridges from breaking as pressure is applied to the center of the cooler.

4.8 Summary

A process has been presented for fabricating in-plane multi-stage thermoelectric coolers. Several designs have been presented, including 1-stage, 5-stage and 6-stage coolers. The devices have been tested, with the best performing 6-stage cooler achieving a ΔT=22.3 K. However, the process used to fabricate these devices is difficult to execute and forces several compromises with respect to the cooler design. Most notably, there is a significant amount of excess oxide at the center stages, which reduces the effectiveness of these stages. Additionally, because the devices are released at the time of deposition, the temperature of the inner stages is not well controlled during deposition, making it difficult to achieve optimum properties across all of the stages. These problems were addressed by development of a new process which is presented in Chapter 5.
Chapter 5
Single Wafer Process

5.1 Introduction

The devices presented in Chapter 4 were capable of achieving temperatures of 22.3 K, but the process ultimately had several drawbacks which limited the manufacturability of the coolers, as well as the performance of the devices produced. The most significant problem, from a performance standpoint, was the low effective ZT of the inner stages, caused by the large amount of oxide spanning them. This limited the performance of those stages to less than 1/10th of the expected intrinsic performance.

Additionally there were significant drawbacks to the process itself. First, handling the 100 μm thick silicon wafers used to fabricate the top layer of the structure was a challenge. The wafers would break easily during routine processing steps, and were not stable until they could be bonded to the structural substrate. Because, these wafers required several processing steps prior to bonding, a significant amount of scrap was produced from this part of the process. Second, the previous process required deposition of the TE material onto coolers that were already released. This meant that during the shadowmask alignment procedure the coolers were already in a very fragile state, and aligning the shadowmask, even with protective spacers resulted in many of the oxide bridges breaking, causing very low yield on individual wafers.

Finally, and perhaps most importantly, the silicon-glass-silicon process was not capable of accepting an external device to be cooled. Any potential device attachment
would need to take place at after the release of the silicon had been removed from he backside of the oxide bridges, and at that point the cooler is too fragile to withstand the pressure required for bonding. This represents a critical flaw in the process. Recall from Chapter 1, that the primary motivation for developing a low-power thermoelectric microcooler is to improve the performance of other low-power microdevices such as gyroscopes, passive IR sensors, and low-power LNAs. Without the ability to interface with such devices, the coolers associated with the previous process have failed to answer the fundamental problem motivating the development of low-power thermoelectric coolers.

To address these shortcomings a new cooler structure and fabrication process were developed. The new process is based around a single 500 µm silicon wafer, and was design to produce a structure with lower parasitic losses, as well as more robust process that produces less scrap. It was also designed to allow for the attachment of arbitrary microdevices prior to release of the cooler structure, bringing it much closer to fulfilling the ultimate goals of the research.

This chapter will begin by discussing the structure of the new cooler devices. It will then present the process used to fabricate the devices, and address some of the challenges associated with this process. Next, a 4-stage cooler with novel power distribution will be presented. Following that, a group of serially driven coolers with 1, 2, 3, and 5 stages will be presented, and their performance will be compared. Finally, the die attachment process will be detailed and characterized.
5.2 Device Overview

Figure 5.1: Cross section of a 2-stage cooler implemented using the structure. The thermal properties of the cooler are implemented with an oxide-metal-oxide stack that is supported by oxide pillars.

A new, simplified process was developed that would allow for deposition of the TE elements and die bonding prior to releasing the structure, while also still meeting the thermal properties required for effective multistage cooling. TE materials were again deposited on top of a supporting layer of silicon dioxide. However, unlike the previous process where the glass tether provided mechanical support to the central region and/or some of the inter-stage regions, the new structure supports these regions between stages with silicon dioxide pillars that are anchored into the silicon below the device. By fabricating the pillars as hollow tubes, the parasitic conduction that they contribute can be kept to a minimum. For example, a pillar that is 150 μm long, with a 15 μm x 15 μm square cross section and an oxide wall thickness of 0.8 μm will have a thermal conductance of 3.5E-7 W/K for a thermal conductivity of 1.1 W/m-K. For comparison, a single TE element with dimensions 3 μm x 100 μm x 30 μm has a thermal conductance of 1E-5. If there were a pillar supporting every single TE element in a one stage cooler,
parasitic thermal conductance due to the pillars would be less than 10% of the total thermal conductance of the device. As Chapter 3 established, this is low enough that it will not have a significant impact on the performance of the device.

The use of the oxide pillars also makes it possible to reduce the parasitic thermal conduction compared to the previous design by reducing the amount of oxide used to support the TE. The devices presented in Chapter 4 had oxide directly underneath the TE, but also bridging the entire gap between the inter-stage regions. Analysis in Chapter 3 showed that this created a significant reduction in the effective ZT at the inner stages, and resulted in a large decrease in performance. In the new structure there is still oxide directly underneath the TE material, but all the oxide between adjacent TE elements can be removed. This is possible because each inter-stage region can be supported with multiple pillars, and the oxide must only support the TE material. This is in contrast to the previous design in which the oxide bridges also had to support one or more of the silicon inter-stage regions. By removing the excess oxide, the parasitic conduction of the first stage of the cooler can be reduced by approximately 50%, and the parasitic conduction of the fifth stage and sixth stages will be reduced by 90% or more.

Multistage cooling also requires providing sufficient thermal conduction in regions between consecutive stages. In the previous designs, this function was performed by bulk single crystal silicon. In the new device, the silicon is replaced by 0.5 µm gold thin-films that are encapsulated in silicon dioxide. The same type of analysis that was used in Chapter 4 to evaluate silicon as the thermal conductor for the previous structure can be used to evaluate the choice of gold in this case. Chapter 3 established that the quantity $K_{inh}\Delta \rho/(\alpha^2T_hN_2)>10$. If we consider a simple cooler with 1 TC at stage 2 and
2 TCs at stage 1, and $\rho=12.9 \ \mu\Omega\cdot\text{m}$, $\alpha=210 \ \mu\text{V/K}$ and a TC aspect ratio of 266,666 (corresponding to 3 $\mu$m x 100 $\mu$m x 80 $\mu$m) then $K_{in}$ should be more than 3.86$\times10^{-5}$. If the gold is 0.5 $\mu$m thick and 100 $\mu$m wide, then the two thermocouples of the first stage can be separated from the thermocouple of the second stage by up to 780 $\mu$m each (See Figure 5.2 for illustration). This indicates that the gold provides enough thermal conductance for a realistic cooler design to be largely unaffected by the substitution of thin gold film for the thick silicon used in the previous design. Having the gold underneath a layer of silicon dioxide decreases the thermal conductance of the path slightly. However, the thermal conductance through 1 $\mu$m of oxide over an area of 150 $\mu$m x 100 $\mu$m is 0.015 W/m-K. This is 3 orders of magnitude greater than the minimum allowed thermal conductance of the gold, and will have a negligible effect on the total thermal conductance in the inter-stage region.

Figure 5.2: Illustration of maximum dimension for inter-stage thermal conduction.

The final requirement for multistage cooling is the need for more increased heat removal at the outer stages. Like the previous designs, this structure allows this requirement to be met by placing additional thermocouples at the outer stages. However, because this structure allows the oxide between TE elements to be removed, there is more flexibility in the layout of the stages. In particular, there is no longer a need to minimize
the perimeter of the inner stages in an effort to minimize parasitic thermal conduction. Instead, the cold stage can be made an arbitrary size without having an adverse affect on the performance of the cooler.

The above structure also deviates from the devices presented in the previous chapters by allowing for the attachment of other devices such as the gyroscopes and passive IR sensors discussed in Chapter 1. To allow for the die attachment, two features were added to the structure. The first feature is a metal bond that is grown by electroplating. The bond ring is separated into 8 segments and provides both the electrical and thermal connections between the cooler and device being cooled. The second feature is the inclusion of the pre-etched recess below the region where the die will sit. This feature gives the cooled device some separation from the hotter wafer, reducing parasitic thermal conduction through the air. It is pre-etched into the wafer and coated with oxide early in the process in order to facilitate faster release at the end of the process, and it will be discussed in more detail in section 5.3.
5.3 Process description

<table>
<thead>
<tr>
<th>Step 0</th>
<th>The process begins with a 500 mm thick silicon wafer.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Deep holes that will later become the pillars are etched into the silicon.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2</th>
<th>A shallower recess is etched below the region where a die will be bonded.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3</th>
<th>A layer of LPCVD oxide is deposited and it forms a conformal coat on the inside of the deep trenches.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Step 4</td>
<td>A layer of metal is deposited on top of the oxide and forms the interstage thermal conductor.</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 5</th>
<th>PECVD oxide is deposited to cover the metal.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 6</th>
<th>The oxide layers are etched in BHF to expose the silicon.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 7</th>
<th>A bond ring is grown by electroplating, which will be used for die attachment.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 8</th>
<th>A layer of metal is deposited and patterned, which will be used for the electrical interconnections.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image9.png" alt="Diagram" /></td>
<td><img src="image10.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Step 9

The TE materials are deposited by shadow masking.

Step 10

A die is attached to the cooler using a gold-indium TLP bond.

Step 11

The entire structure is released with Xenon difluoride.

5.3.1 Process-related challenges and design decisions

Unlike the process presented in Chapter 4, where a number of standard processing steps had to be changed to accommodate the use of thin silicon wafers, this process is generally more straight-forward, with two exceptions. One main challenge is caused by the recess etched below the attached MEMS die. The recess is 100 µm deep, and spinning photoresist would produce very non-uniform layers. This was particularly true on devices near the edge of the wafer, where there would be a significant build up of photoresist on the inner edge of the recess, and significant thinning of resist on the outside part of the recess. The thickest areas were more than 20 µm thick, while the thinnest areas were only 3-4 µm. Such large differences in thickness made lithography
difficult because the thinnest areas would significantly over-develop and delaminate before patterns in the thickest areas completely developed.

If the only factor at stake was the photoresist, the simplest solution would be to move the etch step in order to form the recess later in the process; however, the recess is etched early in the process for a particular reason. If the etch were done as the step directly prior to the TE deposition, there would be no way to deposit oxide on the entire wall of the recessed cavity. In such a case the XeF$_2$ would start etching directly in and down from the bottom of the recess. That would create a situation where the pillars closest to the recess could be undercut and released. Placing the recess etch as the second step in the process and then coating the walls with LPCVD oxide before metal deposition means that the walls are completely coated and etching can only begin at the top surface during the release step. To fix the problem, photoresist was manually dripped into the recesses, and cured without spinning. A second layer of photoresist was then applied and spun on the wafer. This technique allowed for much better photoresist coating to be achieved, and made lithography near the edges of the recesses more repeatable.

5.3.2 Full Process Flow

The following steps detail the fabrication process summarized above.

1. The holes for the oxide pillars are etched.
   1.1. 5 µm of spr220 PR are spun on the surface of a 500 µm thick silicon wafer using the ACS200.
   1.2. The PR is exposed for 12 seconds on the MA-6 contact aligner.
   1.3. The wafer is baked on a hotplate at 115 °C for 90 seconds and developed for 30 seconds in MIF 300 on the ACS200.
1.4. The wafer is etched in the STS Pegasus DRIE tool for 45 minutes using LNF recipe 2 to form the deep holes for the pillars.

1.5. The PR is stripped in PRS2000, rinsed for 3 minutes in DI water and dried in an SRD.

2. The shallow recess DRIE etch is performed.

2.1. AZ9260 PR is spun at 3000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

2.2. The wafer is exposed for 55 seconds on the MA-6 contact aligner.

2.3. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

2.4. The wafer is etched in the STS Pegasus DRIE tool for 12 minutes using LNF recipe 2 to form the shallow recess.

2.5. The PR is stripped in PRS2000, rinsed for 3 minutes in DI water and dried in an SRD.

3. The bottom layer of oxide is deposited.

3.1. Pre-furnace clean is performed on the wafer.

3.2. Low pressure chemical vapor deposition (LPCVD) is used to deposit 0.8 µm of silicon dioxide on the surface of the wafer.

4. The first layer of metal is deposited.

4.1. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.

4.2. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.
4.3. AZ9260 PR is spun at 2000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

4.4. The wafer is exposed for 75 seconds on the MA-6 contact aligner.

4.5. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

4.6. 1000 A of chrome and 5000 A of gold are deposited onto the surface of the wafer using the Enerjet evaporator.

4.7. Lift-off is performed in acetone. An ultrasonic bath is used to speed the lift-off process. The wafers are rinsed with acetone and then with IPA and dried with a nitrogen sprayer.

5. The second layer of oxide is deposited.

5.1. PECVD is used to deposit 0.6-1.0 µm of oxide onto the surface of the wafer. The oxide is deposited with a substrate temperature of 380 °C.

6. The oxide layers are patterned.

6.1. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.

6.2. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.

6.3. 1827 PR is spun at 2000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

6.4. The wafer is exposed for 35 seconds on the MA-6 contact aligner.

6.5. The wafer is developed for 90 seconds in MIF 319, rinsed for 3 minutes in DI water, and dried in an SRD.
6.6. The wafer is etched in BHF for 7 minutes, rinsed in DI water for 3 minutes, and dried in an SRD.

6.7. The PR is stripped from the wafer in PRS2000, and the wafer is rinsed for 3 minutes in DI water and dried in an SRD.

7. The metal bond rings are deposited.

7.1. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.

7.2. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.

7.3. AZ9260 PR is spun at 2000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

7.4. The wafer is exposed for 75 seconds on the MA-6 contact aligner.

7.5. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

7.6. 1000 A of chrome and 5000 A of gold are deposited onto the surface of the wafer using the Enerjet evaporator. This will be the seed layer for electroplating the bond ring.

7.7. Lift-off is performed in acetone. An ultrasonic bath is used to speed the lift-off process. The wafers are rinsed with acetone and then with IPA and dried with a nitrogen sprayer.

7.8. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.
7.9. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.

7.10. AZ9260 PR is spun at 1000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

7.11. The wafer is exposed for 75 seconds on the MA-6 contact aligner.

7.12. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

7.13. 6 µm of gold are electroplated, using the patterned PR as a mold.

7.14. The PR is stripped from the wafer in PRS2000, and the wafer is rinsed for 3 minutes in DI water and dried in an SRD.

7.15. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.

7.16. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.

7.17. AZ9260 PR is spun at 1000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

7.18. The wafer is exposed for 90 seconds on the MA-6 contact aligner.

7.19. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

7.20. The chrome/gold seed layer is etched away, and the wafer is rinsed for 3 minutes in DI water and dried in an SRD.

7.21. The PR is stripped from the wafer in PRS2000, and the wafer is rinsed for 3 minutes in DI water and dried in an SRD.
8. The second layer of interconnect metal is deposited.

8.1. The recesses are filled with 1813 PR and cured for 10 minutes on a 110 °C hotplate.

8.2. HMDS is spun onto the wafer at 4000 rpm for 10 seconds, and the wafer is baked at 110 °C for 90 seconds.

8.3. AZ9260 PR is spun at 2000 rpm for 30 seconds, and baked for 20 minutes in an oven at 90 °C.

8.4. The wafer is exposed for 75 seconds on the MA-6 contact aligner.

8.5. The wafer is developed for 90 seconds in AZ400K, rinsed for 3 minutes in DI water, and dried in an SRD.

8.6. 1000 A of chrome and 5000 A of gold are deposited onto the surface of the wafer using the Enerjet evaporator. This will be the seed layer for electroplating the bond ring.

8.7. Lift-off is performed in acetone. An ultrasonic bath is used to speed the lift-off process. The wafers are rinsed with acetone and then with IPA and dried with a nitrogen sprayer.

9. The TE material is deposited.

9.1. The shadow mask for the Bi$_2$Te$_3$ pattern is aligned to the device wafer and clamped. The wafer is loaded into the TE evaporator and pumped to a pressure of less than 2e-6 Torr. Bi$_2$Te$_3$ is deposited using a flux ration of Te:Bi of 3.0, and substrate temperature of 260 °C. After deposition the wafer is allowed to cool to less than 100 °C before venting the chamber.
9.2. The shadow mask for the $\text{Sb}_2\text{Te}_3$ pattern is aligned to the device wafer and clamped. The wafer is loaded into the TE evaporator and pumped to a pressure of less than $2\times 10^{-6}$ Torr. $\text{Sb}_2\text{Te}_3$ is deposited using a flux ration of $\text{Te}:\text{Sb}$ of 2.4, and substrate temperature of 230 °C. After deposition the wafer is allowed to cool to less than 100 °C before venting the chamber.

10. The Die is attached. See section 5.6 for details

11. The cooler is released.

11.1. The TE is protected with a coat of photoresist, which is sprayed onto the wafer. It is exposed through a special mask which has recesses to accommodate the attached die.

11.2. The wafer is diced or cleaved into groups of up to 4 die.

11.3. The groups of die are released in $\text{XeF}_2$ using the Xactix etching tool. 4 coolers can be released using 200 etching cycles at 30s per cycle with a pressure of 3 Torr.

11.4. An oxygen plasma ash is used to release to completed device.

5.4 The 4 stage device

Using a simplified version of the process described in section 5.3, a 4-stage cooler with a unique current distribution was developed. The coolers described in Chapter 4, and the coolers described by the process in 5.3 rely on using one region of material as the thermal conductor of the inter-stage region (i.e. silicon or the lower metal layer), and a second region of material as the electrical conductor connecting the thermocouples.
However, the design presented in this section combines both functions, electrical and thermal conduction, into a single region of material.

5.4.1 Current drive mechanism

The designs that have been presented so far have all had TCs arranged in series. The electrical conductors are thermally connected to the inter-stage thermal conductor, but are electrically isolated from it, as shown in Figure 5.3(a). This is necessary to maintain the proper current flow through the series TCs. To combine the electrical conductor and the thermal conductor into a single region requires moving away from a driving current through the TCs serial.
Figure 5.3: Illustration of heat and current flow in (a) a conventional, serially-driven 2-stage cooler, and (b) a two stage cooler using a resistive network to distribute current. The electrical schematic of the device shown in (b) is provided as (c).

To understand how this can be accomplished, a diagram of a 2-stage cooler and its equivalent electrical circuit is shown in Figure 5.3 (b-c). In this example, 3 voltage inputs are supplied to the hot side of the first stage, and the current is distributed through the resistive network of the cooler. If the n-type and p-type have the same resistance and
are the same size, the current will be the same through all of the TE legs, resulting in a cooler that has the same cooling characteristics as a 2-stage, serially driven cooler with 3 TCs at stage 1 and 1 TC at stage 2. For a 2-stage cooler, this design approach produces a device that is the equivalent to the design proposed by O’Brien [81]. When extended to 3 or more stages however, the two designs diverge. A 3-stage example of the current design, as well as the O’Brien design, is shown in Figure 5.4 (a-b).

The O’Brien design requires only two inputs to function correctly, but it requires a significant variation in the size of the TE elements at each stage. Although not shown in the schematic, the O’Brien design also requires variations in the size of the TE elements within a stage if the optimal current flow is to be achieved. The design presented here, on the other hand, uses multiple voltage inputs (4 inputs are required for a three stage cooler) but all the TE legs can be made the same size. Uniform size of the TE is an advantage when working with the current process, because the minimum size of the elements is limited by the resolution of the shadow mask deposition process. At the same time, very wide TE legs are impractical because they would take a very long time to undercut with the XeF$_2$ etch process. Long release etches are undesirable because they can result in undercutting the oxide support pillars before the cooler is released.
Figure 5.4: (a) Illustration of a previously devised three stage cooler that distributes current through a resistive network. (b) A 3-stage cooler using the resistive network proposed in this section. (c) A modified version of the cooler shown in (b). It combines inter-stage regions of the same nominal voltage.

The design shown in Figure 5.4(b) is the simplest implementation of this design principle, and is helpful for examining how heat and current flow through the cooler.
However, it is not optimal because it contains isolated regions in the inter-stage zone between stages 1 and 2 which do not have higher cooler capacity at stage 1 compared to stage 2. To fix this problem, inter-stage zones that share a common nominal voltage can be combined into a single zone as shown in Figure 5.4 (c)

5.4.2 Structure and process simplification

Using this design allows the process described in section 5.3 to be simplified. Because the electrical and thermal conductors are combined into a single layer of material, the process no longer requires two metal layers. This means that both Step 5, the PECVD oxide deposition, and Step 8, the second metallization step, can be removed from the process. In addition, the 4 stage devices that were fabricated were not intended to demonstrate the die attachment process. As a result, Step 2, the DRIE etched recess, and Step 7, the bond ring deposition, were also omitted from the fabrication process. The center of the device was modified to include a perforated membrane that supported a resistive temperature sensor for measuring the performance of the devices. The resulting device cross section is shown in Figure 5.5 and can be fabricated with as few as 7 masks.

![Cross section of a 2-stage cooler using the simplified variation of the single wafer process.](image)
5.4.3 Device parameters

A four-stage cooler was designed to make use of this simplified process. It used an equivalent of 30 TCs at the first stage, 14 TCs at the second stage, 6 TCs at the 3rd stage and 2 TCs at the 4th stage. Each TC was 60 μm long by 100 μm wide, and deposited on an oxide bridged that was 120 μm wide. The total size of the cooler was 4.3 mm by 4.3 mm, and the size of the cold stage was 3.1 mm by 3.1 mm. The cold stage was supported by 16 oxide pillars. The inter-stage zone connecting stage 3 and stage 4 was not supported. The zone connecting stage 2 and stage 3 was supported by 6 pillars along each side of the cooler, and the zone connecting stage 1 and stage 2 was supported by 2 pillars along each side of the cooler. The design parameters for this 4-stage device are summarized in Table 5.1, and an SEM view of the entire cooler is shown in Figure 5.6. Figure 5.7 shows an angled view of the corner of the cooler with several oxide pillars clearly visible below the plane of the cooler. The design was not without flaws, however. The lack of supporting oxide pillars at the edges of the inter-stage zone between stages 3 and 4 allowed the edges to curl due to intrinsic film stress, as shown in Figure 5.8. While not always a problem, this curling resulted in breakages of the TE legs in several devices.

Table 5.1: Design parameters of the 4-stage cooler fabricated using resistive network current distribution.

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of TCs</td>
<td>30</td>
<td>14</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Number of pillars at the cold side of the TC</td>
<td>8</td>
<td>24</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>TE Width</td>
<td>120 μm</td>
<td>60 μm</td>
<td>2μm</td>
<td>200μm x 100 μm</td>
</tr>
</tbody>
</table>
Figure 5.6: A top-down SEM view of the 4-stage cooler with resistive network current distribution. The individual stages and the cold platform are highlighted.

Figure 5.7: An SEM showing the oxide pillars supporting the cold platform and stages 1 and 2 of the 4-stage cooler. This image also shows how the oxide has been removed, except directly under the TE materials.
Figure 5.8: An SEM image of the inter-stage zone between stages 2 and 3 of a 4-stage cooler. The zone is not supported by any oxide pillars, and curling of the inter-stage zone is clearly visible. This is in contrast to Figure 5.7, which shows very little deflection of the interstage zones where they are supported by the pillars.

5.4.4 Performance

5.4.4.1 Simulated Cooling Performance

This device was modeled using FEM in Ansys instead of the analytical model described in Chapter 3. This decision was made because the inter-stage zones are separated into multiple segments, and each of the segments between any two given stages can potentially have its own unique temperature (i.e. there are 4 inter-stage segments that connect stage 1 and 2, and each can be at a different temperature). This results in a much more expansive system of equations. In the case of this 4-stage cooler, a system of 50 equations would be necessary to model the cooler. Instead of setting up the system by hand, FEM analysis allows the model to be constructed and graphically depicted. Solving the FEM model is still slower than it would be with the analytical model, particularly because the model needs to be solved over a range of input voltages in order to find the
optimal voltage. In sum, the FEM simulation has an easier setup than Ansys, as well as the ability to take 2-D layout into account, which is very useful in this scenario.

The performance of the cooler was simulated using the optimal material properties that can be achieved by co-evaporated TE materials, as reported in section 2.5.2 of Chapter 2. The simulated cooler also used 6 µm thick thermoelectric films. The 4-stage cooler requires 5 input voltages set at \( +V_1, +V_2 \), and 0, where \( V_1 = 2V_2 \). Figure 5.9 shows a schematic of a 4 stage cooler with 1 TC at the 4\(^{th}\) stage, and the arrangement of the TE elements and input voltages used in the model. This schematic represents half of the fabricated cooler, which has twice as many TCs at each stage. Simulating only half of the cooler will not affect the simulated cooling performance as long as the parasitic thermal conductance due to the pillars is also reduced to half the actual value. This was done by including a parasitic thermal conductance of 1e-5 W/K between the ambient temperature and the cold platform. The simulation was iterated over a range of values for \( V_1 \) (with \( V_2 \) adjusted accordingly) and the optimal voltages were found to be \( V_1 = 90 \) mV and \( V_2 = 45 \) mV, resulting in a minimum simulated temperature of 258 K, at an ambient temperature of 300 K, as shown in Figure 5.10. This is a \( \Delta T \) of 42 K.
5.4.4.2 Measured Performance

The thermal isolation of the cooler was first tested to verify that the oxide pillars provided adequate thermal isolation from the substrate. The cooler was placed in a vacuum probe station, and the resistive temperature sensor was used to apply a thermal...
load, while at the same time measuring temperature. The resulting data is shown in Figure 5.11, and the device was found to have a total thermal resistance of ~42,500 K/W.

![Figure 5.11: Thermal resistance measurement of a 4-stage cooler fabricated using the simplified single wafer process. The thermal resistance is 42,539 K/W.](image)

The released 4-stage coolers were mounted in ceramic dual inline pin (DIP) packages and tested in a vacuum bell jar with electrical feed-throughs to a DIP socket. This technique was used for testing because of the large number of inputs needed to properly test the circuit. The vacuum probe station used in the previous tests only supports six probes. The 4-stage device, however, requires 5 different voltage levels for the cooler itself, plus 4 connections for the resistive temperature sensor. The maximum cooling was found to occur with $V_1=90$ mV, and $V_2=40$ mV, very close to the simulated values, but the cooling was only measured at $\Delta T=17.6$ K. The difference between the measured and simulated results can be explained by the fact that the deposited TE materials did not have the same properties as the optimal films obtained earlier. The
The resistivity of the films on the wafer was 40 µΩ-m for Bi₂Te₃ and 20 µΩ-m for Sb₂Te₃. Additionally, the combined Seebeck coefficient was estimated by thermal load to the cold platform, and by measuring the voltage generated across the device. This gave a result of 320 µV/K, or an average of 160 µV/K per material. When these new values are used in the simulation, along with the thinner 2.2 µm film thickness, the predicted cooling becomes ΔT=18.7 K, just over a degree different from the measured cooling. The results of the new simulation are shown in Figure 5.12. Two additional coolers were measured, and achieved ΔT=12.3 K and ΔT=8.2 K. However, these were only partial coolers. In each case the 4th stage TE for one half of the cooler was broken. This meant that these coolers had only half the heat removal capacity of the full cooler but still had roughly the same level of parasitic thermal conductance due to the oxide pillars. This resulted in the lower performance.

Figure 5.12: FEM simulation of a cooler with resistive network current distribution, using the material measured in section 5.4.4.2.
5.5 1, 2, 3, and 5-Stage coolers

5.5.1 Parameters

The 4-stage coolers presented in the previous section demonstrated a simplified version of both the structure and process presented in sections 5.2 and 5.3. To demonstrate the complete process, new coolers were designed that included 1-stage, 2-stage, 3-stage and 5-stage variations. These coolers once again connect all of their TCs in serial so that only a single current input and a ground connection are needed to operate the cooler. All of the coolers share a common design for the cold platform. It measures 7 mm on aside, and it is supported by 52 oxide pillars. The number of pillars was increased compared to the previous design to provide additional support for attached devices. The coolers also share similar stage designs. Stage 1 is the same for all of the coolers and has a total of 64 TCs, 16 on each side of the cooler. The second stage of all the multi-stage coolers has 32 TCs, and the 3rd stage of the 3-stage and 5-stage coolers has 16 TCs. The 4th and 5th stages of the 5-stage cooler have 8 and 4 TCs respectively. The TE elements are all 80 µm long and 80 µm wide. This represents a tripling of the aspect ratio compared to the 6-stage coolers presented in Chapter 4, and is intended to further reduce power consumption. The design parameters are summarized in Table 5.2, and top down SEMs of all the coolers are shown in Figure 5.13. One notable change is the reduction in contact area compared to the previous designs. This is possible because the higher aspect ratio increases the resistance in the circuit due to the TE material. The important parameter is the ration between parasitic resistance and the resistance of the TE material itself. If the resistance due to the TE material increases, then the parasitic resistance can also increase, without impacting performance.
Table 5.2: Design parameters for the 1, 2, 3 and 5-stage coolers fabricated using the single wafer process.

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of TCs</td>
<td>64</td>
<td>32</td>
<td>16</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Number of pillars at the cold side of the stage</td>
<td>20</td>
<td>20</td>
<td>12</td>
<td>8</td>
<td>52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TE Width</th>
<th>TE Length</th>
<th>TE Contact Area</th>
<th>Oxide Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 µm</td>
<td>80 µm</td>
<td>125 µm x 90 µm</td>
<td>1.4 µm</td>
</tr>
</tbody>
</table>

Figure 5.13: Top down SEM images of 5-stage (annotated), 1-stage (left), 2-stage (center), and 3-stage coolers (right), pictured released but without TE. Only a quarter of each cooler is shown because the devices are too large to visualize completely in the SEM.
5.5.2 Simulated Performance

To understand how different aspects of the design affected the performance of the coolers, all the variations were simulated with three different levels of model completeness. Because all of these devices use TCs that are connected in serial, the 1-dimensional mathematical-based model could again be used, instead of FEM. First the devices were simulated without taking any parasitic thermal conduction into account. This is equivalent to removing any supporting oxide layers, as well as all of the supporting oxide pillars. The optimal material properties found in Chapter 2 were used for the simulations. For Bi$_2$Te$_3$, $\alpha$=208 $\mu$V/K and $\rho$=19 $\mu$\Omega-m. For Sb$_2$Te$_3$, $\alpha$=160 $\mu$V/K and $\rho$=13 $\mu$\Omega-m. For both materials, the thermal conductance was set based on the estimate of 1 W/m-K, calculated from the performance of the 1 and 6-stage coolers presented in Chapter 4. 2 $\mu$m thick films were used for all the simulations. The results of this simulation are shown in Figure 5.14. Under these conditions the simulated 1-stage cooler is able to reach a minimum temperature of 265 K at an input current of 7.5 mA corresponding to 64 mW of power consumption. The simulated 2-stage cooler reaches 257 K at 5.0 mA corresponding to 46 mW of power consumption. The simulated 3-stage cooler reaches 255 K at 4.2 mA corresponding to 39 mW of power consumption, and the simulated 5-stage cooler reaches 264 K at 3.7 mA corresponding to 33 mW of power consumption.
While the new process has substantial parasitic thermal conduction due to a decrease in the oxide spanning the stages, the TE materials still need to be deposited on an oxide film that remains in the final structure. The next set of simulations included the thermal effects of this supporting oxide. The thickness of the simulated oxide layer was 1.7 µm. The results of this simulation are shown in Figure 5.15. Under these conditions the simulated 1-stage cooler is able to reach a minimum temperature of 264 K at an input current of 7.5 mA. The simulated 2-stage cooler reaches 253 K at 5.0 mA. The simulated 3-stage cooler reaches 244 K at 4.1 mA. The simulated 5-stage cooler reaches 228 K at 3.2 mA. The performance of all 4 of the coolers is shown, scaling equally due to the inclusion of the oxide parasitic thermal conductance.
The final set of simulations includes both the support oxide layer below the TE material and the supporting oxide pillars. Each pillar is assumed to have a thermal conductance of $5 \times 10^{-6}$ W/K. This is slightly higher than the calculation given for a 150 µm long pillar earlier in the chapter because the pillars in these designs should be approximately 100 µm long when the devices are fully released. The results of this simulation are shown in Figure 5.16. Under these conditions the simulated 1-stage cooler is able to reach a minimum temperature of 265 K at an input current of 7.5 mA. The simulated 2-stage cooler reaches 257 K at 5.0 mA, corresponding to 46 mW of power consumption. The simulated 3-stage cooler reaches 255 K at 4.2 mA, corresponding to 39 mW of power consumption. The simulated 5-stage cooler reaches 264 K at 3.7 mA, corresponding to 33 mW of power consumption.

The results of these simulations show that while the performance of the 5-stage cooler using this design can be expected to perform better than the 6-stage cooler presented in Chapter 4, it is not the optimal design. While the 1-stage and 2-stage coolers perform approximately the same with and without the inclusion of the oxide pillars in the
simulation, the simulated 3-stage cooler is only able to cool 2 K more than the 2-stage cooler, and 11 K less when simulated without the supporting oxide pillars. The 5-stage cooler only outperforms the 1-stage cooler by 1 K.

The reason that the five stage cooler actually performs worse than the 2-stage and 3-stage devices lies in the fact that all the variations use the same cold platform, supported by 52 pillars. The 1-stage and 2-stage coolers have a large number of TCs directly removing heat from the cold platform, and the thermal conductance between the bulk wafer and the cold platform is dominated by the TE and the oxide supporting the TE. On the other hand, the 3-stage and the 5-stage devices have much lower thermal conductance due to the TE and supporting oxide, and the thermal conductance due to the oxide pillars is still the same. As a result, the ratio between the TE thermal conductance and the parasitic thermal conductance, $K_{TE}$ and $K_{par}$, can become significant, reducing the overall performance of the cooler. In the case of the 5-stage device, the thermal conductance of the of 4 TCs at the 5th stage is approximately $8E-6$ W/K, while the conductance due to the oxide pillars is $2.5E-5$ W/K, over three times the thermal conductance of the TE across the fifth stage.
5.5.3 Fabrication issues

The devices fabricated using the full process, with two layers of metal and two layers of oxide, exhibited a failure mode that was not observed in the simpler fabrication process used for the 4-stage coolers reported above. When devices are completely released there is widespread cracking of the Sb$_2$Te$_3$ and the oxide below it, as shown in Figure 5.17. This creates open circuits and renders the cooler in-operable. The breakage appears to occur only after the removal of the protective photoresist films, and occurs where the TE covers the bottom metal layer. This has been observed in 100% of the devices from which the photoresist is removed, and on devices from several different wafers. The fact that the breakage only occurs on the Sb$_2$Te$_3$ legs indicates a stress problem. Tests of the intrinsic film stress for both Bi$_2$Te$_3$ and Sb$_2$Te$_3$ show that Bi$_2$Te$_3$ has a tensile stress of 65 MPa, while Sb$_2$Te$_3$ has a tensile stress of 128 MPa, further pointing to stress as the culprit, since the Sb$_2$Te$_3$ has the higher stress of the two films and is also the point of failure. It is not clear why the devices fabricated using the complete
process are exhibiting this problem, while the devices fabricated using the simplified process did not.

![Crack in TE and Oxide](image)

**Figure 5.17:** An SEM image of an antimony telluride leg of a TC. The crack in the antimony telluride is clearly visible.

### 5.5.4 Measured Performance and Material Analysis.

A 3-stage cooler with the photoresist intact was tested in the vacuum probe station. To measure the temperature of the cold platform, the bond ring was replaced with a resistive temperature sensor operated in a 4-point measurement configuration. The TE films were deposited in the Lesker co-evaporator, and this wafer represented the first coolers to receive films from this tool. The resistivity of the thin films was measured on the wafer and found to be 50 µΩ-m for Bi₂Te₃ and 18 µΩ-m for Sb₂Te₃. This is much higher than the optimal films, and has a direct impact on the performance that can be achieved. The total resistance of the cooler was 3800 Ω, and it achieved cooling of 3 K at an input current of 1 mA. Such a small optimal current also indicates that the Seebeck coefficient of these materials is also very far from the optimal point. For a 3-stage cooler
with the given resistance to have the optimal current at 1 mA, the average Seebeck coefficient needs to be 80 \( \mu \text{V/K} \), indicating that the films in this case were far from optimal.

### 5.5.5 Revised performance predictions

Based on the material property data estimated from the 3-stage cooler, all 4 variations were simulated again, and the results are shown in Figure 5.18. Even with such poor material properties, the simulated performances of the coolers exceed the measured performance. The 3-stage cooler is predicted to have a \( \Delta T = 5.3 \text{ K} \) with an input current of 0.95 mA. The difference between this simulated value for the 3-stage cooler is likely due to the 7-10 \( \mu \text{m} \) layer of photoresist which remained on the device during testing. The 1-stage and 5-stage coolers are both predicted to cool by 4 K, and the 2-stage is predicted to cool by 5 K. This analysis makes it clear that in order to achieve any meaningful performance from the revised design, better TE material properties must be obtained from the new deposition tool.
Figure 5.18: Performance predictions based on the material properties extrapolated from the measured performance of a 3-stage cooler.

5.6 Die attachment

As discussed in the introduction to this Chapter, die attachment is a critical consideration in the development of a practical microcooler. For the cooler to be widely applicable it must be capable of integrating with a wide variety of device, these devices may, in general, be produced using a variety of different fabrication process. While the size, shape and performance of the cooler can be tailored to the specific application, the die attachment process should be design to require the smallest possible changes to the process of the target device. This ensures that preparing the device for integration with the cooler will not result in a change in its baseline performance.

There are several die attachment scenarios to consider when design the cooler for integration with devices. The first is that of a die which does not need to interface with the environment. This would include devices such as LNAs, other electronics and gyroscopes. In this case flip-chip bonding can be used. The materials needed for the
bonding process can be deposited directly onto the existing pads of the target device as additional step at the end of its fabrication process. The devices can then be cingulated and bonded using a template or bonded at the wafer level.

The second scenario involves devices such as infrared sensors and imagers that need to receive a signal, e.g. IR or visible light, that would be blocked by the bulk silicon of the cooler wafer. One solution to correctly expose the sensor is to mount the target device facing upward. In this case making an electrical connection between the active area of the targeted device and the cold platform would require through wafer via to be integrated into the targeted device. Such a feature would be strongly dependent on the process used to fabricate the device. A second approach is to extend the recess in the substrate of the cooler through the wafer, providing a window. This approach would allow for flip chip bonding to again be used, requiring minimal changes to the process for the targeted device, but it would require a more complex package with, both a top and a bottom cap.

The final scenario involves attaching silicon based MEMS devices to the cooler. These devices are at risk of being attacked by the XeF$_2$ during the release process. To protect the devices some form encapsulation require during the XeF$_2$ etch process. This encapsulation could potentially take a number of different forms, and the ideal solution would vary depending on the type of device being integrated with the cooler. The simplest of option would be to coat any exposed silicon with a protective layer of metal or dielectric that would protect it from the etchant. For devices where such a coating would be impossible or impractical two other more complex options exist.
The first is to hermetically seal the device with a low profile cap, such as the thin film packages described in [98] prior to bonding, and allow the cap the remain as part of the finally product. A second option is to encapsulate the cooler polymer such as photo-resist during the release process, and the remove the polymer by O$_2$ plasma ashing. However, this would require access to the front side of the targeted device. That would again mean integrating through-wafer vias into the process of the targeted device or providing an opening from the backside of the wafer to the targeted device to allow for application and removal of the polymer.

Each of these scenarios presents its own challenges, many of which would be unique to the specific device being integrated with the cooler. However, despite the differences all three scenarios need a bonding process that can provide a thermal and an electrical connection between the cold platform and the device being cooled. The remainder of this section will discuss an appropriate bonding technology to be used with the thermoelectric microcoolers.

### 5.6.1 TLP bonding

For the thermoelectric coolers to be useful, there must be a means of attaching a target device to the cold platform. The connection between the target device and the cold platform must provide both electrical and thermal conduction. For this reason, bonds using solders or inter-metallic compounds are appealing. An additional constraint of the die attachment process is that it should be low temperature. The die attachment is performed after the deposition of the TE material, and heating the material to temperatures greater than its deposition temperature can change the composition and the properties of the TE films. Low temperature solder bonding and transient liquid phase
(TLP) bonding were previously explored at the University of Michigan using In/Au, Ni/Sn, and Au/Sn [99],[100]. TLP bonding works by placing a low melting point material such as indium or tin between two higher melting point materials, as illustrated in Figure 5.19. When the stack is heated, the middle layer will melt and diffuse into the adjacent metal layers, forming an intermetallic compound that is solid. Based on the previous work, an Au/In bonding process was chosen because it can be performed at temperatures as low as 200 C, and because the intermetallic diffusion is slower than Au/Sn, resulting in a strong homogenous bond.

To implement the bond, 5-6 µm thick bond pads were electroplated on the cold platform of the cooler, on top of a seed of 0.1 µm Cr and 0.5 µm Au. The bond pads on the target device consist of a stack of 0.1 µm of NiCr for an adhesion layer, 0.5 µm of Au, 2 µm of In, and finally an additional 0.1 µm of Au to prevent oxidation. The bonding is performed at 200 C, but the temperature is allowed to briefly reach as high as 230 C to facilitate rapid and uniform melting of the In layer.

![Figure 5.19: Cross section illustration of the TLP bonding process showing the low melting temperature metal as purple and the high melting temperature metal in yellow. The illustration includes (a) The bond pads prior to bonding, (b) the arrangement at the initiation of the bond, and (c) the diffused intermetallic compound after completing the bonding process.](image-url)
5.6.2 Target device alignment and transfer

Previous work with die transfer has used a template wafer to guide pre-cingulated devices into place. This technique has proven effective, but can result in die that are misaligned by up to 30 µm. Unfortunately, this particular process cannot tolerate such large misalignments. In order to release the cooler in a reasonable amount of time and not undercut the supporting pillars, the cold platform requires perforations which are very close to the attached device, but not covered by it. Bonding an un-diced wafer of target devices onto the cooler wafer can provide alignment to within 10 µm, but presents the problem of how to remove the kerf of the device wafer. Dicing can be used, but risks damaging the TE material already deposited on the wafer. Instead, the device die are partially cingulated using DRIE, but are held to the bulk silicon wafers by 4 silicon tethers, one at each corner. The tethers allow the entire wafer of device die to be aligned to the cooler wafer with good precision. After bonding, the tethers are broken and the bulk of the wafer is lifted away, leaving only the device die attached to the cooler wafer.

5.6.3 Proof of Concept

To demonstrate the die attachment using Au/In TLP bonding, target devices that contained a resistive temperature sensor were fabricated. Each die contained 8 bond pads, arranged around the perimeter of the die. The perimeter was notched in order to decrease the amount of undercutting required to release the cooler, while still distributing the weight of the die directly onto the oxide pillars. The wafers were aligned in the Suss BA-6, and the bond was performed in the Suss Sb6e. Figure 5.20 shows a die attached to an unreleased cooler. Figure 5.21 shows the bond part of the bonded area for an attached die that has been pried away. The silicon of the attached die has fractured and remains
attached to the cooler, indicating that the strength of the bond was stronger than the fracture strength of silicon.

![SEM image of a die bonded to an unreleased cooler structure.](image1)

**Figure 5.20:** An SEM image of a die bonded to an unreleased cooler structure.

![SEM images at different magnification levels of the bond region after the attached die has been pried away. Silicon has fractured, indicating a strong bond.](image2)

**Figure 5.21:** SEM images at different magnification levels of the bond region after the attached die has been pried away. Silicon has fractured, indicating a strong bond.
5.7 Packaging

Chapter 4 demonstrated the importance of operating the coolers at vacuum in order to reduce parasitic thermal conduction through the air. Methods for bonding a hermetically sealed package capable of maintaining vacuum include anodic bonding [95], Au/Si eutectic bonding [101], and TLP bonding [99],[100],[102]. As with the die attachment process, a Au/In TLP bond is appropriate for this application because it can be performed at low temperatures. To achieve the lowest possible pressure, the wafers need to dwell in a high vacuum environment for up to 24 hours before bonding. This allows for outgassing of adsorbed substances from the surface of the cooler and package. Such a long outgassing step is necessary because the wafers cannot be heated to speed the outgassing process, without prematurely activating the indium. After bonding, a getter must be included in the package design. Once bonding is completed under vacuum, the getter is activated by heating the entire device. The materials in the getter react with the gasses remaining in the cavity to reduce the pressure further. Combined with the 24 hour dwell time, this process has produced packaged devices with pressures as low as 20 mTorr [103]. The getter is typically activated at temperatures in excess of 300 °C. However, it has been demonstrated that activation can also be achieved using temperatures as low as 200 C, but requires a longer dwell time at the activation temperature [103].

The vacuum packaging process also needs to include electrical feed throughs. This could most easily be achieved with an under bump metallization process, where metal feedthroughs pass underneath the capping bond ring and are isolated from the cap by a dielectric passivation layer. During the bonding step the liquid indium flows around
bumps created by the feedthroughs and completely seals the package. It has previous been shown that for the indium to completely seal around a 0.5 µm thick feedthrough, it must be at least 0.88 µm thick (assuming a heating rate during bonding of 60 °C) [103]. The process of vacuum packaging has not been implemented as part of this thesis, but the approach presented above could be appended to the existing process with no major changes, except for the addition of a bond ring for the cap.

5.8 Summary

This chapter presented a process for fabricating a cooler that could be implemented using a single wafer. The new process used a metal thin film as the thermal conductor instead of silicon, and it allowed for the coolers to be released after deposition of the TE material. A 4-stage cooler with a new current distribution scheme was fabricated using a simplified version of the process, and it achieved 17.6 K of cooling. 1, 2, 3, and 5-stage coolers were fabricated using the full fabrication process, but due to breakage of the Sb₂Te₃, only the one 3-stage device could be tested. Simulation showed that the 2-stage and 3-stage devices had the potential for cooling in excess of 40 K, but poor material properties prevented this from being realized. Die attachment was demonstrated using TLP bonding, and a path for incorporating vacuum packaging into the process was presented.
Chapter 6

Conclusion and Future Work

Thermoelectric micro-coolers hold the potential to improve the performance of a wide range of devices, including passive IR sensors, gyroscopes and other resonant MEMS devices, and low-noise analog circuitry. This thesis has investigated several different designs for thermoelectric micro-coolers using two different processes. In addition it presented data related to the deposition of thermoelectric thin films by co-evaporation, and a model for predicting the performance of thermoelectric coolers. The model was also used to determine design parameters for multistage thermoelectric coolers. An overview of the results and future work associated with this project are presented in the following sections.

6.1 Conclusion

Bismuth telluride and antimony telluride were the TE materials used for this project. They were deposited through a co-evaporation process that was optimized for ZT by varying the flux ratio of the constituent elements during deposition, and the temperature of the substrate. Bismuth telluride was produced with a Seebeck coefficient of 208 $\mu$V/K, a resistivity of 19 $\mu\Omega$-m, and a ZT of 0.40. Antimony telluride was produced with a Seebeck coefficient of 160 $\mu$V/K, a resistivity of 13 $\mu\Omega$-m, and a ZT of 0.34. A newer tool was also used to deposit thermoelectric thin films, but is still being characterized, and has not yet matched the performance of the original films.
A 1-dimensional analytical model was implemented using Mathematica to predict the performance of new cooler designs and estimate the material properties in devices being tested. The model was also used to understand how parameters such as aspect ratio and relative levels of parasitic thermal and electric conduction influence the performance of the cooler. Specifically, parasitic electrical resistance per TC should be less than 10% of the resistance of the TC, and parasitic thermal conductance per TC should be less than 10% of the conductance of the TC. In addition, a metric was introduced to evaluate the thermal conductance of the inter-stage zone for particular designs.

Two different processes were developed for fabricating thermoelectric micro-coolers. The first process involved a silicon-glass-silicon stack of three wafers. The primary features of the cooler were implemented on the top silicon wafer, while the glass wafer was patterned to provide mechanical support to the fragile structure above it. Four variations of coolers were implemented using this process. An un-optimized 5-stage cooler was implemented first, and was able to achieve 8 K of cooling. A six-stage cooler was then designed and fabricated, which achieved 16.6 K of cooling. A modified version of this cooler was able to achieve 22.3 K of cooling using only 24.7 mW of power.

The second process utilized only a single wafer, and was developed to solve several of problems related to the silicon glass silicon process. The single-wafer process used metals for both the thermal conductors and the electrical conductors, and it provided mechanical support to the cold platform and inter-stage regions using oxide pillars. As a result, the intra-stage oxide bridges no longer play a structural role beyond supporting the TE material, allowing for the removal of any interstage oxide that is not directly supporting the TE material. By doing this the parasitic thermal conduction at the center
stages was greatly reduced. Additionally, the single wafer processes was designed to release the TE coolers only after the deposition of the TE material. This provided a stable subtracted to process on until releasing with XeF$_2$ in the final step of fabrication, thus reducing the number of devices lost breakage early in the fabrication process. By placing the release at the end of the process, the single-wafer process also accommodates the attachment of arbitrary devices for cooling, such as gyroscopes or other sensors,

Simulations suggested that cooling of 45 K using only 41 mW of power could be achieved if the optimal material properties could be paired with a 3-stage device. 1, 2, 3 and 5-stage coolers were fabricated, but only a 3-stage device has been tested. The 3-stage device showed only 3 K of cooling, and indicated that the thermoelectric material properties on the wafer were far below the best values that have been achieved. A 4-stage device with a novel current drive mechanism was also fabricated using a simplified version of this process, and it achieved cooling of 17.6 K. A method for die attachment using Au/In TLP bond was also proposed, and proof of the concept was demonstrated.

Although they have not yet yielded high performance coolers, the 1 and 3-stage devices based on the second process continue to hold the greatest potential going forward. They offer reasonable cooling levels and low power consumption. The 1-stage device will be best for applications that present a higher heat load, while the 3-stage device provides an additional 10 K of cooling while still maintaining low power consumption.
6.2 Future work

6.2.1 Thermoelectric Materials

The key to developing a reliable thermoelectric micro-cooler is consistently producing high quality thermoelectric materials. The best materials produced to date at the University of Michigan have $ZT<0.5$ at room temperature, and the observed $ZT$ properties in several fabricated devices, is lower yet do to increased electrical resistance. At the same, bulk materials have long demonstrated $ZT$ approaching 1 at 300 K, and other researchers in thin film deposition of TE materials have demonstrated films with $ZT$ exceeding 0.7. Such increases $ZT$ are critical to improved performance and manufacturability of planar micro TE coolers. By applying higher quality materials to the same structures that have been demonstrated in this thesis, higher temperature differentials can be achieved. At the same time, higher yields could be achieved by targeting similar temperature differentials, and increasing the robustness of the structure.

Work should continue on optimizing the materials and understanding the deposition parameters that affect the quality and repeatability of the TE thin films. To achieve higher levels of $ZT$ and reduce process drift, more complex ternary and quaternary films should be explored. An investigation of co-evaporated superlattice TE films should also be undertaken to see if the high $ZT$ achieved with these materials can be reproduced using co-evaporation techniques.

Finally a detailed investigation into the effects of substrated and contact material, on the TE thin film should be conducted. The current processes use silicon dioxide and gold because small scale tests suggested that these were the best materials for obtaining good contact resistivity, adhesion and $ZT$. However, a larger scale investigation,
encompassing a greater range of materials should be conducted to verify the optimal materials for both the supporting dielectric and the electrical contact region.

6.2.2 Structure and Process

The failure mechanism associated with the breakage of the antimony telluride legs of the new devices should be further investigated. It will be important to understand why this occurred when the full process was implemented, but not on the simpler version of the process. The solution may be as simple as providing a sturdier platform for the TE by using a thicker layer of oxide, or it may require a larger change to the process to compensate for the intrinsic stress in the films.

The die attachment process needs to be pursued further. The single-wafer process was designed to accommodate die attachment, but more characterization is necessary. First the minimum level of support for the cold platform needs to be investigated. As shown in Chapter 5, the presence of the oxide pillars decreases the overall performance of the device compared to an ideal cooler. It also affects the optimum number of stages that should be employed by the cooler to achieve the highest temperature difference. Maximizing temperature performance will rely on minimizing the number of pillars, while still providing adequate mechanical support.

A second challenge that requires additional investigation is the technique to protect sensitive devices during the XeF\(_2\) etch release process. This process was discussed in Chapter 5 and multiple possible solutions were discussed. This included temporary encapsulation with a polymer, or permanent encapsulation using a low profile packaging technique such as thin film encapsulation. The next step is to demonstrate
implementation of one or more of these techniques in combination with an actual silicon MEMS device.

Finally, low temperature vacuum packaging should be pursued further. Although low temperature bonding has been demonstrated in the past, there is more work that can be done in this area. Pressures below 100 mTorr should be achieved to minimize the effect of parasitic thermal conduction through the air. Although pressures as low as 20 mTorr had previously been demonstrated there was still significant variance in between the devices in that study, and more work should be done to improve the vacuum level uniformity of the process.

6.3 Projected Long Term Results

If this project were to be pursued for an additional several years, with a focus in the areas of materials and structure as outlined in Section 6.2, it is anticipated that devices capable of consistently generating a ΔT=50 K, could be consistently achieved. Such a device would most like have 2 or 3 stages, and its power consumption, and heat removal capability would be determined by aspect ratio of the TCs. It is not expected that performance could be achieved to rival the best bulk, or even micro scale multistage coolers, largely because of the need for a supportive membrane below the TE material, which reduces ZT. However, low power devices capable of reaching temperatures of 250 K are of interest for a number of applications such as gyroscopes, passive IR sensors and micro-chromatographs, as described in Chapter 1.
References


[38] G.E. Bulman, E. Siivola, B. Shen, and R. Venkatasubramanian, “Large external


[87] Y. Cheng and C. Shih, “Maximizing the cooling capacity and COP of two-stage


[92] “PECVD Silicon Dioxide.”

[93] “summaryofprop.pdf.”

[94] “Parylene Properties & Characteristics.”


